QUELQUES ÉLÉMENTS THÉORIQUES AUTOUR DU LEXIQUE

INTRODUCTION

1) Quelques distinctions autour du mot LEXIQUE :

-SYNTAXE/LEXIQUE : la syntaxe est faite de l'ensemble des règles permettant de combiner les mots pour constituer des syntagmes et les syntagmes pour constituer des phrases ; le lexique regroupe l'ensemble des unités pouvant être ainsi combinées, soit **les mots** de la langue.

-LEXIQUE/VOCABULAIRE : l'ensemble des mots disponibles dans une langue donnée constitue son **lexique**. L'ensemble des mots dont dispose effectivement un locuteur de cette langue constitue son **vocabulaire**.

2) La notion de signe linguistique établie par Saussure en 1916 est nécessaire pour étudier le lexique. Le signe, c'est l'association d'un signifiant (Sa) – image acoustique ou visuelle – et d'un signifié (Sé) – contenu sémantique, concept. À côté, le référent, c'est ce que désigne le signe : une réalité.

3) L'un des outils de base pour étudier le lexique est le dictionnaire. Le maniement et l'étude du dictionnaire commencent dès le CE1. Les premières leçons de vocabulaire y sont consacrées. Ce travail porte essentiellement sur le dictionnaire de langue, à distinguer du dictionnaire encyclopédique.

-le dictionnaire encyclopédique ou dictionnaire de « choses » : il donne en priorité des informations sur le référent, c'est-à-dire la réalité désignée par le mot. Il est souvent accompagné d'illustrations. Ex : *Petit Larousse illustré.*

-le dictionnaire de langue ou dictionnaire de « mots » : donne en priorité des informations sur le signe linguistique et notamment ses relations structurales avec d'autres éléments du vocabulaire. Les deux plus connus sont le *Robert* (grand ou petit) et le *Lexis* chez Larousse.

4) D'où viennent les mots du français ?

Les origines de la langue : tous les mots que nous employons ont une histoire. La plupart des mots du français sont issus d'un étymon latin. Ainsi, l'étymon du français *maître* est *magister*. Certains adjectifs de couleur sont d'origine germanique : lors des invasions, la langue des envahisseurs a supplanté les correspondants latins. C'est le cas pour les adjectifs suivants : *bleu*, *blanc*, *gris*, *brun*. Le français standard a aussi recueilli certains termes des dialectes d'oïl ou d'oc qu'il a pourtant supplantés : *amour*, *abeille* viennent de l'occitan ; *congère*, *moraine* du franco-provençal.

Les emprunts aux langues étrangères : ils rendent compte de l'influence de l'histoire, de la géographie, des échanges économiques ou culturels.

Le dictionnaire de référence en la matière est le *Dictionnaire historique de la langue française (*A.

Rey, chez Robert). Ex : *tohu-bohu*, mot emprunté à l'hébreu ; des mots entrés très récemment dans le dictionnaire : *surimi* (Japon), *taliban* (Afghanistan) ; une adaptation au système phonologique et graphique du français est parfois nécessaire : l'anglais *packet-boat* donne *paquebot* ; *riding-coat* donne *redingote*.

La création lexicale : selon les besoins, on forme de nouveaux mots à partir de monèmes existants.

I- LA MORPHOLOGIE LEXICALE

Elle s'intéresse à l'organisation formelle du lexique, la forme des diverses unités lexicales ainsi que les procédés de formation utilisés pour produire ces unités. La morphologie analyse donc la structure interne des unités, leur construction, c'est-à-dire ses constituants. Ces derniers sont appelés **morphèmes**. Un morphème est une unité formelle minimale, indécomposable et porteuse de sens, obtenue par segmentation des énoncés. Donc l'unité de l'analyse morphologique ne sera pas le mot (sauf lorsqu'il est simple, c'est-à-dire indécomposable : *jupe*, *grain*, *table*). La morphologie lexicale prend pour unité le morphème lexical dit aussi lexème et n'étudie que les mots construits ou complexes en mettant en évidence les relations de forme et de sens qui unissent les unités lexicales. [La morphologie lexicale est à distinguer de la morphologie grammaticale qui s'occupe des morphèmes grammaticaux ou morphogrammes (marques du nombre, genre, personne, temps).]
Ex : ***injustement*** : 3 morphèmes, où chaque morphème est l'association d'une forme et d'un sens :

Forme :	***in-***		***juste***		***-ment***
		+	« conforme à	+	
Sens :	« non »		la justice »		« de manière »

Il existe deux grands modes de création lexicale : la dérivation et la composition. Mais il faut relever aussi : la conversion, l'abréviation, la siglaison.

A- LA DÉRIVATION : ce procédé crée une nouvelle unité en combinant une base et un ou des affixes. (dérivation = base + affixe). [Cette notion de dérivation renvoie à ce qui est intitulé « famille de mots » dans les programmes. Une famille de mots, c'est l'ensemble des mots dérivés, par suffixe(s) ou préfixe(s), d'un même mot-souche.]

-la base : est le morphème lexical qui porte le signifié (Sé), qui a une valeur dénominative. Ex : *chant-* dans *chanter*, *créd-* dans *incrédule*. La base, c'est ce qui reste d'un mot dérivé si on lui enlève

tous ses affixes.

-l'affixe : préfixe ou suffixe. Il n'a pas de valeur dénominative, il possède une signification ou valeur.

.le **préfixe**, placé à gauche de la base, modifie rarement la classe grammaticale de celle-ci. Sa fonction est essentiellement sémantique. Ex : *re-* ou *ré-* peut marquer la réitération : recommencer, reprendre, rouvrir...

.le **suffixe**, postposé à la base, modifie le plus souvent et le sens et la catégorie grammaticale de celle-ci. Il est un indicateur de classe grammaticale. Ex : le suffixe *-age* sert à former des noms. C'est un suffixe nominal. Il signifie « action de » dans *nettoyage, repassage* ; mais aussi « état de » dans *concubinage, esclavage*. Un suffixe peut apporter d'autres informations grammaticales. Ex : *-ise, -ance, -tion* créent des substantifs féminins ; *-age et -isme* des substantifs masculins.

À partir de bases généralement nominales, verbales et adjectivales, la suffixation produit des noms, des verbes, des adjectifs et des adverbes. Ex : *pauvre* donne *pauvreté* : [[-]Adj+Suf]N ; *arroser* donne *arrosoir* : [[-]V+Suf]N (cas de nominalisation)

Attention : l'adjonction d'un suffixe peut modifier la base (méchant/méchamment) ; parfois, un phonème/graphème de cohésion peut être indispensable : dur-*c*-ir.

Remarques générales :

-Un mot dérivé peut être formé de la combinaison de plusieurs préfixes et suffixes : anti-constitution (n) elle-ment.

-Certains affixes présentent des variantes de forme, phoniques et ou graphiques, conditionnées par l'environnement phonologique et morphologique de l'affixe. Ex : les préfixes in-/im-/ir-/il-/ ; re-/ré-/r. Les suffixes -ité/-té ; -tion/-ation/-ition.

-La **formation parasynthétique** désigne l'adjonction simultanée d'un préfixe et d'un suffixe (embourgeoiser) sachant que le préfixé et le suffixé correspondants n'existent pas : *embourgeois/ *bourgeoiser.

B- LA CONVERSION : ou « transfert », ou encore « dérivation impropre » : elle consiste en un changement de classe grammaticale (et donc de sens) sans changement de forme. « Le rouge lui va bien » : le nom rouge est obtenu par simple recatégorisation grammaticale de l'adjectif rouge.

.Beaucoup de conversions en nom à partir d'un/d'une

-adjectif : le vrai, le faux

-pronom : le ça, le moi

-verbe : le boire, le manger, un militant

-préposition : un avant, un contre

-adverbe : le pourquoi et le comment
-conjonction : avec des si et des mais

.Conversions en adjectif de

-nom : une veste *marron*
-adverbe : un homme *bien*

.Conversion en adverbe de

-adjectif : parler *fort*, chanter *faux*
-préposition, par effacement du GN introduit : « je suis *pour* ».

C- LA COMPOSITION : deux éléments ou plus, appelés bases, sont associés (composition = base + base). Ces éléments sont en principe des unités lexicales autonomes (contre-poids). Mais ils peuvent ne pas avoir d'existence autonome (camé-scope ; écho dans écho-graphie).

-Les noms composés sont nombreux : grande surface, mode d'emploi, touche-à-tout.
-mais aussi les verbes : court-circuiter
-les adjectifs : bleu clair, bon dernier
-les déterminants : une tonne de, une dizaine de, beaucoup de, vingt-quatre
-les adverbes : ci-devant, ci-dessus
-les pronoms : celui-ci, le mien, lequel

Pour les déterminants, les adjectifs et les adverbes composés, le caractère plurigraphique des unités fait qu'on les désigne aussi par le terme de « locution ».

1) La composition proprement dite ou composition populaire regroupe tous les mots composés dont les éléments sont des mots français qui ont une existence autonome par ailleurs : *portefeuille, porte-monnaie, chaise longue*. Les éléments réunis dans un mot composé forment une unité de sens nouvelle dont la signification dépasse celle de ses éléments pris isolément : une chaise longue n'est pas littéralement une chaise qui est longue mais globalement, un fauteuil pliable destiné au repos en position allongée.

L'interprétation des mots composés français dépend d'abord de la nature (catégorielle et sémantique) des constituants et de leur ordre qui, dans les formes non savantes, place l'élément recteur (déterminé) avant l'élément régi (déterminant). C'est l'élément recteur qui donne le genre du mot composé. Ex : N1 – N2 : un *timbre-poste,* un *bébé-éprouvette,* une *cigarette-filtre*.

Comme la dérivation, la composition n'est plus perçue dans de nombreux mots considérés aujourd'hui comme des termes simples : *maintenir*

Concernant les modes d'association possibles entre les constituants d'un composé, trois cas se rencontrent :

-soudure graphique des constituants : *portefeuille, clairvoyant*. Ils tendent à être perçus comme des mots simples, surtout lorsque la soudure modifie la prononciation : *vinaigre*.
-les éléments sont liés par un trait d'union : *chef-lieu, sourd-muet, cerf-volant* (rappel : la réforme de l'orthographe de 1990, dans une volonté de simplification, préconise de remplacer le trait d'union par la soudure pour les composés en : contre-, intra-, entr(e)-, extra-, infra-, ultra-.
-les éléments sont séparés par un blanc graphique : *bon sens, eau de vie, garde champêtre, à cause de, afin que*.

2) La composition savante : la plupart des termes scientifiques et techniques du français sont formés d'éléments savants grecs et latins. On parle dans ce cas de composition savante. Les composés savants sont formés par la réunion de bases grecques ou latines, qui ne constituent généralement pas des unités lexicales autonomes (sauf abréviations : *télé, auto*)

Certains composés savants sont empruntés en bloc au latin ou au grec : aqueduc : *aquaeductus* ; amphibie : *amphibios*. Mais la plupart des composés ont été forgés en français ; cependant, la voyelle de transition -o- relie en principe deux éléments de composition grecs (ex : *baromètre*) et la voyelle -i- deux éléments latins (ex : *digitigrade*). Ces formations comportent le plus souvent des éléments de même origine.
-composés grecs : anthropologie, thalassothérapie, démocratie, topographie.
-composés latins : apiculture, multicolore, homicide, viticole
-parfois on trouve des composés hybrides : *génocide* : *genos* signifie « la race » en grec, *caedere* signifie « tuer » en latin ; ou base française et élément grec : hypotension, archiplein.

La composition savante peut être associée à la dérivation affixale. Ainsi, beaucoup de noms savants fonctionnent comme des bases pourvues du suffixe nominal -ie comme élément final : *agronomie, radioscopie, télépathie.*

Par ailleurs, certains éléments savants, par leur productivité et leurs combinaisons, se comportent comme des préfixes et tendent à être assimilés à des préfixes : anti- : antidote, antigel, hypo- : hypoallergénique, hypotension. Des éléments adverbiaux comme bien-, mal-, non-, arrière- (bienheureux, malformé, arrière-plan) ou prépositionnels comme après-, avant-, contre-, entre-, sur-, sous, outre-, sans- (service après-vente, avant-coureur, contre-épreuve, entre-deux, surdétermination, sous-louer, outrepasser).

D- SIGLES ET ABRÉVIATIONS : ces deux procédés contribuent à l'extension du vocabulaire, limitent l'emploi de mots longs et économisent ainsi la dépense articulatoire et mémorielle.

1) L'abréviation constitue une réduction du signifiant d'un mot, le signifié restant

inchangé. Elle donne lieu à la troncation des mots longs (+ de trois syllabes), par retranchement d'une ou plusieurs syllabes (le découpage syllabique n'est pas toujours respecté : *manif, prof, perm, clim* se terminent par la première consonne de la syllabe suivante).
-**l'aphérèse** tronque l'avant d'un mot : (auto) *car*
-**l'apocope**, la plus fréquente, tronque la fin d'un mot : *télé* (vision), *prof* (esseur)

Le procédé de l'abréviation est très productif dans la langue contemporaine et les mots abrégés servent de base pour la formation de mots nouveaux : *bus*, *abribus*.

Les abréviations affectent souvent des composés savants particulièrement fréquents, ou des termes du parler scolaire ou militaire : *stylo, bac*... Cet termes abrégés, familiers au départ, ont éclipsé aujourd'hui le terme complet : *cinématographe, stylographe*.

2) La siglaison : les sigles sont des unités formées par la suite des lettres initiales de mots composés. À l'écrit, ils se reconnaissent à l'emploi des lettres capitales, éventuellement séparées par des points. Leur importance dans la langue contemporaine va croissant. Ils désignent des organisations politiques : PS, MODEM, LCR ; syndicales : FO, SNES ; étatiques : USA ; internationales : ONU, UNESCO... Chaque domaine d'activité possède ses sigles : ESPE, RASED, MGEN pour le milieu enseignant.
Pour la prononciation des sigles : on donne le nom de chaque lettre comme dans la récitation de l'alphabet (ex : le PS) ; mais de plus en plus, le sigle est prononcé comme un mot ordinaire, quand il peut se lire syllabiquement (ex : RASED).
Certains sigles st devenus des noms véritables (ex : *radar* : radio detecting and ranging). Dans ces cas, l'usage des majuscules n'est plus nécessaire.
Certains sigles servent de base à la formation de mots dérivés par suffixation : *énarque, capésienne, bédéphile*.

II- LA SÉMANTIQUE LEXICALE

Nous l'aborderons sous deux angles :
-la différenciation entre les plans de la langue et du discours
-les relations sémantiques entre unités lexicales

A- LE PLAN DE LA LANGUE ET DU DISCOURS [comme on distingue langue (système de signes) et langage (la langue en contexte, mise en pratique)]. Ne pas confondre avec le discours qui s'oppose au récit.

1) sens/signification : en principe, le terme de « sens » renvoie à ce que signifie le mot ou la

locution en discours, c'est-à-dire dans une situation donnée : en contexte, tant à l'oral qu'à l'écrit. « Signification », c'est ce que signifie un mot ou une locution en langue (définition du dictionnaire). (mais cette terminologie n'est pas stable, certains linguistes inversent).

2) dénotation/connotation : ces termes concernent aussi bien le plan de la langue que celui du discours.

-la dénotation désigne ce que, en vertu de sa signification, le mot ou la locution représente, ce à quoi il ou elle permet de référer (potentiellement sur le plan de la langue, effectivement sur le plan du discours). La dénotation est pour cela parfois considérée comme la **signification de base du mot.** Ex : *la faux* dénote un instrument formé d'une lame arquée fixée au bout d'un long manche dont on se sert pour couper le fourrage.

-la connotation renvoie, elle, en rapport avec le référent potentiel ou virtuel, possible autrement dit (plan de la langue) ou avec tel référent précis dans une situation donnée (plan du discours), à une ou à des **valeurs sémantiques ressenties comme secondes et particulières.** Les connotations sont socialement et culturellement marquées, elles véhiculent des représentations ou idées. Ex : *la faux* connote la violence, le temps qui passe, la mort ; elles véhiculent des symboliques : le blanc, par exemple, connote la virginité. Les connotations peuvent aussi relever d'un niveau de langue déterminé.

B- LES RELATIONS SEMANTIQUES DANS LE LEXIQUE : les mots se définissent sémantiquement en partie les uns par rapport aux autres. Ces relations sont dites structurelles ou structurales. Les linguistes parlent de « structuration du lexique » et étudient les relations de synonymie, antonymie, hyperonymie, hyponymie. La polysémie est répertoriée parfois au titre de ces relations : une relation sémantique interne à un même mot ou locution. L'homonymie, la paronymie.

1) La polysémie : les différentes significations ou acceptions d'un même mot : différentes significations se rattachent à une même forme.

Ex : un *canard* c'est l'animal, le sucre trempé, la fausse note, le journal.

Il existe un lien entre ces différentes significations :

-la métaphore : similarité de posture : le canard plonge dans l'eau comme le sucre dans le café.

-la métonymie : une relation de contiguïté fait qu'un même mot désigne un nouvel objet : le contenant pour le contenu : boire un *verre* ; un instrument pour son utilisateur : le premier *violon*

-la synecdoque : un rapport de contiguïté et une relation logique d'inclusion expliquent le lien. Ex : la partie pour le tout : un *toit* pour un logement.

L'ensemble des différentes acceptions d'un même mot forme son **champ sémantique**.

Il faut rattacher à la notion de polysémie, celle de sens propre ou premier (le plus usuel) et sens second figuré, c'est-à-dire utilisant une image, une analogie pour exprimer une idée. Ce point est important pour expliquer des locutions figées ou expressions lexicalisées : *avoir un cheveu sur la langue*, *avoir un papa gâteau*.

Il faut souligner le caractère massif de la polysémie dans la langue française où plus de 40 % des mots sont plysémiques (c'est-à-dire avec avec au moins deux subdivisions dans un article de dictionnaire).

2) La synonymie : c'est une relation de ressemblance sémantique entre des unités lexicales de même classe grammaticale. La synonymie absolue est rare. Elle est limitée à des contextes.

Ex : un gros homme : corpulent, fort, bedonnant

une grosse faute : grave, sérieuse

un gros mot : grossier, vulgaire

Il faut prendre en compte aussi la question des niveaux de langue, qui modifient le point de vue : il est différent de dire « tu fais la tête » et « tu fais la tronche ».

3) L'antonymie : relation d'opposition entre des unités de même catégorie grammaticale relativement à un même domaine.

On distingue 3 types d'opposition, selon le rapport des unités.

-rapport de réciprocité : il s'agit souvent de verbes : acheter/vendre ; donner/recevoir : le même procès est envisagé selon deux points de vue complémentaires.

-rapport de gradation : deux unités pivots constituent les points de référence aux extrémités d'une échelle d'une propriété susceptible de degrés : froid/frais/tiède/chaud.

-rapport d'exclusion : la négation d'une unité implique l'affirmation de l'autre, sans degré intermédiaire : mort/vivant, vrai/faux, recto/verso.

On différencie aussi :

-les antonymes lexicaux : *bête/intelligent*

-les antonymes morphologiques formés à partir d'un préfixe : *lisible/illisible*

4) L'hypéronymie et l'hyponymie : la relation est une relation hiérarchique d'inclusion. Elle concerne surtout la classe grammaticale des noms.

Ex : *fruit* est l'hyperonyme ; prune, pomme, banane sont des hyponymes.

L'hyperonyme ou terme générique (ou encore « mot-étiquette ») est plus riche référentiellement (renvoie à de nombreuses réalités) mais plus pauvre sémantiquement que l'hyponyme ou terme spécifique. Cette relation est importante dans les travaux de taxinomie (classification d'éléments).

5) L'homonymie : des homonymes sont des termes ayant le même signifiant (la

même forme graphique ou sonore) mais un signifié (sens) différent.
-si l'orthographe est différente, il s'agit d'homophones.
-si l'orthographe est la même, il s'agit d'homophones particuliers, les homographes.
Il est souvent difficile de dire si l'on a affaire à un mot unique polysémique ou à plusieurs mots homonymes. Ex : *voler* (dérober) et *voler* (dans les airs) ont une origine commune, tous deux appartiennent au vocabulaire de la fauconnerie. Les lexicographes (ceux qui élaborent les dictionnaires) ont alors deux possibilités, correspondant à deux philosophies :
-ou bien ils privilégient le point de vue diachronique (historique), l'évolution sémantique au cours des siècles. Ils choisiront alors de faire de *voler* un seul mot polysémique (une seule entrée dans le dictionnaire, donc).
-ou bien ils privilégient le point de vue synchronique, l'état de la langue à un moment donné, et constatent la disparité des sens dans l'usage actuel : ils optent alors pour deux mots, homonymes (deux entrées dans le dictionnaire, donc)
Ex : le mot *langue* : dans le *Petit Robert*, une seule entrée et trois grandes définitions sont proposées. Le terme est donc considéré comme polysémique. Dans le *Lexis* : 3 entrées sont proposées (organe, système de signe, chose en forme de langue). Les trois mots sont considérés comme homonymes.

6) La paronymie : des paronymes sont des mots ayant des signifiants (sonorités) non pas identiques mais proches, leurs signifiés n'ayant aucun rapport. Ils font l'objet de nombreuses confusions et nécessitent donc des mises en garde :
collision/collusion
perpétrer/perpétuer

Ouvrage de référence :

M. Riegel, J.-C. Pellat, R. Rioul, *Grammaire méthodique du français* (1994), PUF, coll. « Quadrige », 2005.

Cécile Yapaudjian-Labat